RECOMBINANT BIOTIN CARBOXYLASE DOMAINS FOR IDENTIFICATION OF ACETYL CoA CARBOXYLASE INHIBITORS

Tedd D. Elich, Sandra L. Volrath, and Stephanie C. Weatherly

Related Applications

This application claims priority under 35 U.S.C. 119(e) from United States Provisional Patent Application 60/401,170, filed August 5, 2002, the disclosure of which is incorporated by reference herein in its entirety.

Field of the Invention

The present invention relates to a peptide comprising a biotin carboxylase domain and fragments thereof useful for the identification of Acetyl CoA carboxylase inhibitors, which in turn are useful among other things as fungicides, insecticides, nematicides,herbicides and pharmaceuticals.

Background of the Invention

Acetyl CoA carboxylase (ACCase) catalyzes the first committed step in fatty acid biosynthesis and has also been chemically validated as an herbicide and fungicide target. Structurally, ACCases are biotinylated, multifunctional enzymes comprised of three domains: a biotin carboxylase domain, a biotin binding site, and a carboxytransferase domain. In prokaryotic ACCases, as well as in the plastidic isoforms of most plant ACCases, the three domains reside on three distinct, dissociable proteins. In contrast, in most eukaryotic ACCases the three domains reside on a single polypeptide of 160 kD to 280kD. In their native state, the eukaryotic enzymes are typically dimers or tetramers ranging in size from approximately 400 – 800 kD.

The ACCase reaction takes place at two catalytic sites via two partial reactions: the ATP dependent carboxylation of the enzyme-bound biotin prosthetic group, and the subsequent transfer of the carboxyl group from biotin to acetyl CoA to form malonyl CoA. The natural product soraphen has been demonstrated to be a broad-spectrum fungicide that acts by inhibiting the biotin carboxylase reaction of

ACCase. ACCase's are known to be low abundant and labile proteins. These properties impede the identification of new ACCase inhibitors.

The present invention provides a peptide comprising a biotin carboxylase domain and fragments thereof useful for the identification of Acetyl CoA carboxylase inhibitors, which in turn are useful among other things as fungicides, insecticides, nematicides,herbicides and pharmaceuticals.

Summary of the Invention

According to embodiments of the present invention, the present invention relates to a peptide comprising an Acetyl CoA carboxylase (ACCase) having a deleted biotin binding domain, having a deleted carboxy transferase domain, and having a functional biotin carboxylase (BC) domain (e.g., capable of binding soraphen). In some embodiments of the invention where such BC domains are used as counterselection agents in conjunction with peptides or BC domains as described above, the peptide or BC domain is non-functional.

"Functional" as used herein refers to a BC domain that binds soraphen with similar affinity as enzymatically active, full length ACCase protein. Thus, "non-functional" as used herein refers to a BC domain that does not bind soraphen. These non-functional BC domains would be functional with respect to enzyme activity/catalytic function when incorporated into an intact ACCase.

The carboxylase (and corresponding peptide) may be from any suitable source, including plant, animal (*e.g.*, mammalian), insect, yeast, and fungal carboxylases/peptides.

According to other embodiments of the present invention, the carboxylase (and corresponding peptide) is from *Ustilago maydis* carboxylase.

According to still other embodiments of the present invention, the carboxylase (and corresponding peptide) is from *Phytophthora infestans* carboxylase.

According to still other embodiments of the present invention, the carboxylase (and corresponding peptide) are from *Magnaporthe grisea, Saccaromyces cerevisiae* and *Homo sapiens*.

According to other embodiments of the present invention, the present invention relates to the molecules described above wherein the respective peptides are each an Acetyl CoA carboxylase (ACCase) having a deleted biotin binding domain,

having a deleted carboxy transferase domain, and having a functional biotin carboxylase domain comprising amino acids as detailed in **SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, or 16**, and functional fragments thereof.

According to other embodiments of the present invention, the molecules described above are each a monomer.

According to still other embodiments of the present invention, the present invention relates to the molecules described above wherein the respective carboxylase domains bind to compounds that modulate Acetyl CoA carboxylase activity.

According to other embodiments of the present invention, the carboxylase domains bind to competitive inhibitors, noncompetitive inhibitors, and also binds to soraphen.

According to other embodiments of the present invention, the present invention relates to a nucleic acid that encodes a peptide comprising an Acetyl CoA carboxylase (ACCase) having a deleted biotin binding domain, having a deleted carboxy transferase domain, and having a functional biotin carboxylase domain, such as described above and further hereinbelow.

According to other embodiments of the present invention, the present invention relates to a recombinant host cell that contains a nucleic acid as described above and expresses the encoded peptide.

According to other embodiments of the present invention, the present invention relates to a method of identifying Acetyl CoA carboxylase inhibitors, or activators, comprising a) combining a peptide as described above and a compound to be tested for the ability to bind to said biotin carboxylase domain, under conditions that permit binding to said biotin carboxylase domain, and b) determining whether or not said compound binds to said biotin carboxylase domain, the presence of binding indicating said compound is or may be an Acetyl CoA carboxylase inhibitor. Such compounds are candidates for and useful as pesticides, including but not limited to insecticides, nematocides, fungicides, and/or herbicides, and/or also pharmaceuticals, including but not limited to antifungals.

According to other embodiments of the present invention, the present invention relates to a method of identifying Acetyl CoA carboxylase inhibitors, further comprising the steps of c) employing a compound identified as binding in step (b) in an assay to detect inhibition of Acetyl CoA carboxylase activity; and d)

selecting a compound identified in step (c) that inhibits Acetyl CoA carboxylase activity.

According to still other embodiments of the present invention, the present invention relates to a method of identifying fungicides, comprising a) combining a peptide as described above and a compound to be tested for the ability to bind to said biotin carboxylase domain, under conditions that permit binding to said biotin carboxylase domain, b) determining whether or not said compound binds to said biotin carboxylase domain, the presence of binding indicating said compound is or may be a fungicide, c) employing a compound identified as binding in step (b) in an assay to detect inhibition of Acetyl CoA carboxylase activity, and d) selecting a compound identified in step (c) that inhibits Acetyl CoA carboxylase activity.

According to still other embodiments of the present invention, the present invention relates to the use of a peptide or compound as described above for carrying out a method as described above.

Brief Description of the Drawings

Figure 1 illustrates full-length ACCase protein from *Ustilago* maydis (pCS11) with the three functional domains detailed.

Figure 2 illustrates soraphen binding to and inhibition of the full-length pCS11 protein.

Figure 3 illustrates soraphen binding to the *Ustilago* BC domain (pCS8) with comparable affinity to full length ACCase (pCS11).

Figure 4 illustrates soraphen binding to *Phytopthora infestans* BC domain.

Figure 5 illustrates the biotin carboxylase domain of Ustilago peptide (pCS8) compared to full-length Ustilago ACCase.

Figure 6 illustrates the amino acid sequence of *Ustilago maydis* ACCase BC domain, amino acids 2–560 (pCS8, **SEQ ID NO: 2**)(Taken from Full Length Amino Acid Sequence for *Ustilago maydis* ACCase, Accession Number: Z46886;A. Bailey, J. Keon, J. Owen, and J. Hargreaves, ACC1 gene, encoding acetyl-CoA carboxylase, is essential for growth in *Ustilago maydis*, *Mol. Gen. Genet.* 249 (2), 191-201 (1995)).

Figure 7 illustrates the amino acid sequence of *Phytopthora infestans* ACCase BC domain, amino acids 1-555 (pCS15, **SEQ ID NO: 4).**

Figure 8 illustrates anion Exchange Chromatography of pCS8 peptide.

Figure 9A illustrates spectrophotometric assay absorbance traces for *E. coli* BC and Figure 9B shows *(i)* spectrophotometric and *(ii)* 14C isotope exchange activity assays on pCS11 protein.

Figure 10 illustrates the alignment of *Ustilago* (**SEQ ID NO: 2**), *Phytophthora* (**SEQ ID NO: 4**), *Magnaporthe* (**SEQ ID NO: 6**) and yeast (**SEQ ID NO: 8**) ACCase BC domains.

Figure 11 illustrates soraphen binding to the *Magnaporthe* BC domain.

Figure 12 illustrates soraphen binding to the human ACC1 BC domain.

Figure 13 illustrates the alignment of the human ACC1 BC (**SEQ ID NO: 10**) and ACC2 BC (**SEQ ID NO: 12**) domains with the *Ustilago* ACCase BC (**SEQ ID NO: 2**) domain.

Figure 14 illustrates dissociation experiments using [^{3}H]-soraphen to determine the soraphen off rate for *Ustilago* ACCase BC domain.

Figure 15 illustrates the binding of [^{3}H]-soraphen A and soraphen C conjugates to (A) the *Ustilago* ACCase BC domain and (B) the full-length *Ustilago* ACCase protein.

Figure 16 depicts soraphen binding to wild type and mutant *S. cerevisiae* ACCase BC domain peptides.

Figure 17 illustrates soraphen A binding to wild-type and mutant full-length *S. cerevisiae* ACCase.

Detailed Description of the Preferred Embodiments

The present invention will now be described more fully hereinafter with reference to the accompanying figures, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. The disclosures of all United States patent references cited herein are to be incorporated by reference herein in their entirety.

Described herein is the use of recombinant, isolated, biotin carboxylase domains for the discovery of new Acetyl CoA carboxylase (ACCase) inhibitors. A biotin carboxylase (BC) domain from the ACCase gene of the basidiomycete *Ustilago*

maydis was isolated, cloned, expressed, and characterized. The isolated BC domain was shown to have similar high-affinity, soraphen-binding properties as the full-length protein. In contrast to the full-length protein (**Figure 1**), however, the BC domain is significantly smaller and can be expressed at higher levels, is more stable, and exists as a monomer. The isolated BC domain is useful for screening new ACCase inhibitors. The BC domain from the oomycete *Phytophthora infestans* was also cloned. A full-length ACCase sequence from this organism has not been published. The appropriate fragment was cloned utilizing PCR using primers derived from published EST's that showed homology to sequences flanking the soraphen-binding domain that was identified in the *Ustilago* gene. The recombinantly expressed *Phytopthora* BC domain exhibited high-affinity soraphen-binding. BC domains from *M. grisea*, *S. cerevisae*, and *H. sapiens* were also similarly cloned and determined to exhibit high-affinity soraphen-binding, thus demonstrating the applicability of this approach to distantly related organisms.

The terminology used in the description of the invention herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention. As used in the description of the invention and the appended claims, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise.

As used herein, an "isolated" nucleic acid (*e.g.*, an "isolated DNA" or an "isolated genomic RNA") means a nucleic acid separated or substantially free from at least some of the other components of the naturally occurring organism or virus, for example, the cell or viral structural components or other polypeptides or nucleic acids commonly found associated with the nucleic acid.

Likewise, an "isolated" polypeptide means a polypeptide that is separated or substantially free from at least some of the other components of the naturally occurring organism or virus, for example, the cell or viral structural components or other polypeptides or nucleic acids commonly found associated with the polypeptide. As used herein, the terms "polypeptide" and "peptide" have the same meaning.

As used herein, the terms "deleted" or "deletion" mean either total deletion of the specified segment or the deletion of a sufficient portion of the specified segment to render the segment inoperative or nonfunctional (*e.g.*, does not encode a functional

peptide, wherein functional is defined as the ability to bind soraphen), in accordance with common usage. *See, e.g.* U.S. Patent No. 6,180,362; U.S. Patent No. 5,689,039.

As used herein, the term modulation of Acetyl CoA carboxylase activity refers to the ability of a compound to alter the activity of the enzyme. The alteration may be by enhancing or decreasing the activity of the enzyme, or by causing the enzyme to function in a manner other than that observed in the absence of the compound.

Also as used herein, the term activator refers to the ability of a compound to initiate and/or enhance Acetyl CoA carboxylase activity (*e.g.*, an agonist).

The term inhibitor as used herein refers to the ability of a compound to decrease and/or terminate Acetyl CoA carboxylase activity (*e.g.*, an antagonist).

As used herein, test compounds refer to compounds that may bind the biotin carboxylase domain, under conditions that permit binding to the biotin carboxylase domain. The presence of binding indicating the compound is or may be an Acetyl CoA carboxylase inhibitor. Moreover, binding of the compound to the biotin carboxylase domain may indicate that the compound may be a fungicide, insecticide, nematicide, or herbicide or may be a pharmaceutical (*e.g.*, a compound that reduces, controls, inhibits or otherwise regulates weight gain in a human or animal subject, particularly compounds that are inhibitors of human or mammalian ACC2, and more particularly compounds that preferentially inhibit or antagonize human or mammalian ACC2 and not human or mammalian ACC1, *see, e.g.*, L. Abu-Elheiga et al., *Science* **291,** 2613 (30 March 2001)). Additionally, binding of the test compound refers to specific binding wherein the binding interaction between the BC domain and test compounds is high. The dissociation constant of the BC domain complexes is from about 10^{-4} M to about 10^{-14} M, more preferably 10^{-7} to 10^{-14}, and still more preferably 10^{-8} to 10^{-14} and most preferably lower than 2×10^{-9} M. The test compound may be identified by any available means, including but not limited to the Evolutionary Chemistry process described herein below.

Amino acid sequences disclosed herein are presented in the amino to carboxy direction, from left to right. The amino and carboxy groups are not presented in the sequence. Nucleotide sequences are presented herein by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented herein in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission, or (for amino acids) by three-letter code, in accordance with 37 C.F.R

§1.822 and established usage. *See, e.g.*, Patent In User Manual, 99-102 (Nov. 1990) (U.S. Patent and Trademark Office).

In general, the term "peptide" refers to a molecular chain of amino acids with a biological activity (e.g., capacity to bind soraphen). If required, it can be modified in vivo and/or in vitro, for example by glycosylation, myristoylation, amidation, carboxylation or phosphorylation; thus inter alia oligopeptides and polypeptides are included. It is understood however that the peptides of the present invention do not extend to native proteins which may contain the disclosed peptides. The peptides disclosed herein may be obtained, for example, by synthetic or recombinant techniques known in the art. It will also be understood that amino acid and nucleic acid sequences may include or exclude additional residues, such as additions or deletions of N- or C-terminal amino acids or 5' or 3' sequences, and yet still be essentially as set forth as disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological activity (e.g., capacity to bind soraphen). Thus, up to about 10, 20, 30, or about 40 amino acids may be deleted from either, or both, the N- and/or C- terminus of the peptide, so long as a functional biotin carboxylase domain (e.g., soraphen binding) is retained. Examples of such peptides are peptides having the amino acid sequence given in **SEQ ID NO: 14, 16 and 17** through **71** herein. Note that, for BC domains of human ACC1 and ACC2, up to 102 and 244 amino acids, respectively may be deleted from the N-terminal end, alone or in combination with the above listed C-terminal deletions, so long as a function biotin carboxylase domain (e.g., soraphen binding) is retained.

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Suitable nucleic acid sequences encoding an ACCase biotin carboxylase (BC) domain (that is, an ACCase having a deleted biotin binding domain and a deleted carboxy transferase domain) include, for example, a nucleic acid encoding a *Ustilago maydis* BC domain. Examples of such are given as **SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, and 15** herein.

Polynucleotides of the present invention include those coding for peptides homologous to, and having essentially the same biological properties as, the peptides disclosed herein. For example, the DNA sequences disclosed herein as **SEQ ID**

NOS:1, 3, 5, 7, 9, 11, 13, and 15. This definition is intended to encompass natural allelic sequences thereof. Thus, isolated DNA or cloned genes of the present invention can be of any species of origin. Thus, polynucleotides that hybridize to any one or more of the DNA sequences disclosed herein as **SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15** and which code on expression for an ACCase BC domain, are also an aspect of the invention. Conditions which will permit other polynucleotides that code on expression for a protein or peptide of the present invention to hybridize to the DNA of **SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15** herein can be determined in accordance with known techniques.

For example, hybridization of such sequences may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions (*e.g.*, conditions represented by a wash stringency of 35-40% Formamide with 5x Denhardt's solution, 0.5% SDS and 1x SSPE at 37°C; conditions represented by a wash stringency of 40-45% Formamide with 5x Denhardt's solution, 0.5% SDS, and 1x SSPE at 42°C; and conditions represented by a wash stringency of 50% Formamide with 5x Denhardt's solution, 0.5% SDS and 1x SSPE at 42°C, respectively) to DNA of **SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15** herein in a standard hybridization assay. *See, e.g.*, J. Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2d Ed. 1989) (Cold Spring Harbor Laboratory). In general, sequences which code for proteins or peptides of the present invention and which hybridize to the DNA of **SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15**, for example, will be at least 60% or 75% identical or homologous, 85% identical or homologous, 90% identical or homologous and even 95% identical or homologous, or more with one or more of **SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, and 15**.

Mathematical algorithms can be used to determine the percent identity of two sequences. Non-limiting examples of mathematical algorithms are the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA 87*:2264, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA 90*:5873-5877; the algorithm of Myers and Miller (1988) *CABIOS 4*:11-17; the local homology algorithm of Smith *et al.* (1981) *Adv. Appl. Math. 2*:482; the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol. 48*:443-453; and the search-for-similarity-method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA 85*:2444-2448.

Various computer implementations based on these mathematical algorithms have been designed to enable the determination of sequence identity. The BLAST programs of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403 are based on the algorithm of Karlin and Altschul (1990) *supra*. Searches to obtain nucleotide sequences that are homologous to nucleotide sequences of the present invention can be performed with the BLASTN program, score = 100, wordlength = 12. To obtain amino acid sequences homologous to sequences encoding a protein or polypeptide of the current invention, the BLASTX program may be used, score = 50, wordlength = 3. Gapped alignments may be obtained by using Gapped BLAST as described in Altschul *et al.* (1997) *Nucleic Acids Res. 25*:3389. To detect distant relationships between molecules, PSI-BLAST can be used. *See* Altschul *et al.* (1997) *supra*. For all of the BLAST programs, the default parameters of the respective programs can be used.

Further, polynucleotides that code for proteins or peptides of the present invention, or polynucleotides that hybridize to that as **SEQ ID NO:1, 3, 5, 7, 9, 11, 13, and 15,** or polynucleotides having sequence identity or homology thereto as described above, for example, but which differ in codon sequence therefrom due to the degeneracy of the genetic code, are also an aspect of this invention. The degeneracy of the genetic code, which allows different nucleic acid sequences to code for the same protein or peptide, is well known in the literature. *See, e.g.*, U.S. Patent No. 4,757,006 to Toole et al. at Col. 2, Table 1.

The production of cloned genes, recombinant DNA, vectors, transformed host cells, proteins and protein fragments by genetic engineering is well known. *See, e.g.*, U.S. Patent No. 4,761,371 to Bell et al. at Col. 6 line 3 to Col. 9 line 65; U.S. Patent No. 4,877,729 to Clark et al. at Col. 4 line 38 to Col. 7 line 6; U.S. Patent No. 4,912,038 to Schilling at Col. 3 line 26 to Col. 14 line 12; and U.S. Patent No. 4,879,224 to Wallner at Col. 6 line 8 to Col. 8 line 59.

PCR is the polymerase chain reaction--a technique for copying the complementary strands of a target DNA molecule simultaneously for a series of cycles until the desired amount is obtained. First, primers are synthesized that have nucleotide sequences complementary to the DNA that flanks the target region. The DNA is heated to separate the complementary strands and then cooled to let the primers bind to the flanking sequences. A heat-stable DNA polymerase is added, and the reaction is allowed to proceed for a series of replication cycles. Twenty will yield

a millionfold amplification; thirty cycles will yield an amplification factor of one billion.

A vector is a replicable DNA construct. Vectors are used herein either to amplify DNA encoding the proteins or peptides of the present invention or to express the proteins or peptides of the present invention. An expression vector is a replicable DNA construct in which a DNA sequence encoding the proteins of the present invention is operably linked to suitable control sequences capable of effecting the expression of proteins or peptides of the present invention in a suitable host. The need for such control sequences will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. Amplification vectors do not require expression control domains. All that is needed is the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants.

Vectors comprise plasmids, viruses (*e.g.*, adenovirus, cytomegalovirus), phage, retroviruses and integratable DNA fragments (*i.e.*, fragments integratable into the host genome by recombination). The vector replicates and functions independently of the host genome, or may, in some instances, integrate into the genome itself. Expression vectors should contain a promoter and RNA binding sites which are operably linked to the gene to be expressed and are operable in the host organism.

DNA regions are operably linked or operably associated when they are functionally related to each other. For example, a promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation. Generally, operably linked means contiguous and, in the case of leader sequences, contiguous and in reading phase.

Transformed host cells are cells which have been transformed or transfected with vectors containing DNA coding for proteins or peptides of the present invention and need not express protein or peptide. However, in the present invention, the cells preferably express the protein or peptide.

Suitable host cells include prokaryotes, yeast cells, or higher eukaryotic organism cells. Prokaryote host cells include gram negative or gram positive organisms, for example *Escherichia coli* (*E. coli*) or *Bacilli*. Higher eukaryotic cells include established cell lines of mammalian origin as described below. Exemplary host cells are *E. coli* W3110 (ATCC 27,325), *E. coli B*, *E. coli* X1776 (ATCC 31,537), E. coli 294 (ATCC 31,446). A broad variety of suitable prokaryotic and microbial vectors are available. *E. coli* is typically transformed using pBR322. *See* Bolivar et al., *Gene* **2**, 95 (1977). Promoters most commonly used in recombinant microbial expression vectors include the beta-lactamase (penicillinase) and lactose promoter systems (Chang et al., *Nature* **275**, 615 (1978); and Goeddel et al., *Nature* **281**, 544 (1979), a tryptophan (trp) promoter system (Goeddel et al., *Nucleic Acids Res.* **8**, 4057 (1980) and EPO App. Publ. No. 36,776) and the tac promoter (H. De Boer et al., *Proc. Natl. Acad. Sci. USA* **80**, 21 (1983). The promoter and Shine-Dalgarno sequence (for prokaryotic host expression) are operably linked to the DNA of the present invention, *i.e.*, they are positioned so as to promote transcription of the messenger RNA from the DNA.

Expression vectors should contain a promoter which is recognized by the host organism. This generally means a promoter obtained from the intended host. Promoters most commonly used in recombinant microbial expression vectors include the beta-lactamase (penicillinase) and lactose promoter systems (Chang et al., *Nature* **275**, 615 (1978); and Goeddel et al., *Nature* **281**, 544 (1979), a tryptophan (trp) promoter system (Goeddel et al., *Nucleic Acids Res.* **8**, 4057 (1980) and EPO App. Publ. No. 36,776) and the tac promoter (H. De Boer et al., *Proc. Natl. Acad. Sci. USA* **80**, 21 (1983). While these are commonly used, other microbial promoters are suitable. Details concerning nucleotide sequences of many have been published, enabling a skilled worker to operably ligate them to DNA encoding the protein in plasmid or viral vectors (Siebenlist et al., *Cell* **20**, 269 (1980). The promoter and Shine-Dalgarno sequence (for prokaryotic host expression) are operably linked to the DNA encoding the desired protein, *i.e.*, they are positioned so as to promote transcription of the protein messenger RNA from the DNA.

Eukaryotic microbes such as yeast cultures may be transformed with suitable protein-encoding vectors. *See e.g.*, U.S. Patent No. 4,745,057. *Saccharomyces cerevisiae* is the most commonly used among lower eukaryotic host microorganisms,

although a number of other strains are commonly available. Yeast vectors may contain an origin of replication from the 2 micron yeast plasmid or an autonomously replicating sequence (ARS), a promoter, DNA encoding the desired protein, sequences for polyadenylation and transcription termination, and a selection gene. An exemplary plasmid is YRp7, (Stinchcomb et al., *Nature* **282**, 39 (1979); Kingsman et al., *Gene* **7**, 141 (1979); Tschemper et al., *Gene* **10**, 157 (1980). This plasmid contains the trp1 gene, which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1 (Jones, *Genetics* **85**, 12 (1977). The presence of the trp1 lesion in the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for metallothionein, 3-phospho-glycerate kinase (Hitzeman et al., *J. Biol. Chem.* **255**, 2073 (1980) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Reg.* 7, 149 (1968); and Holland et al., *Biochemistry* **17**, 4900 (1978), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., EPO Publn. No. 73,657.

Cultures of cells derived from multicellular organisms are a desirable host for recombinant protein synthesis. In principal, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture, including insect cells. Propagation of such cells in cell culture has become a routine procedure. *See* Tissue Culture, Academic Press, Kruse and Patterson, editors (1973). Examples of useful host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, and WI138, BHK, COS-7, CV, and MDCK cell lines. Expression vectors for such cells ordinarily include (if necessary) an origin of replication, a promoter located upstream from the gene to be expressed, along with a ribosome binding site, RNA splice site (if intron-containing genomic DNA is used), a polyadenylation site, and a transcriptional termination sequence.

The transcriptional and translational control sequences in expression vectors to be used in transforming vertebrate cells are often provided by viral sources. For

example, commonly used promoters are derived from polyoma, Adenovirus 2, and Simian Virus 40 (SV40). *See, e.g.*, U.S. Patent No. 4,599,308. The early and late promoters are useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. *See* Fiers et al., *Nature* **273**, 113 (1978). Further, the protein promoter, control and/or signal sequences, may also be used, provided such control sequences are compatible with the host cell chosen.

An origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral source (*e.g.* Polyoma, Adenovirus, VSV, or BPV), or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient.

Host cells such as insect cells (*e.g.*, cultured *Spodoptera frugiperda* cells) and expression vectors such as the baculorivus expression vector (*e.g.*, vectors derived from *Autographa californica* MNPV, Trichoplusia ni MNPV, *Rachiplusia ou* MNPV, or *Galleria ou* MNPV) may be employed to make proteins useful in carrying out the present invention, as described in U.S. Patents Nos. 4,745,051 and 4,879,236 to Smith et al. In general, a baculovirus expression vector comprises a baculovirus genome containing the gene to be expressed inserted into the polyhedrin gene at a position ranging from the polyhedrin transcriptional start signal to the ATG start site and under the transcriptional control of a baculovirus polyhedrin promoter.

Host cells transformed with nucleotide sequences encoding a protein or peptide of the invention may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode a protein or peptide of the invention may be designed to contain signal sequences which direct secretion of the protein or peptide through a prokaryotic or eukaryotic cell membrane. Other constructions may be used to join sequences encoding the protein or peptide to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin,

and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the protein or peptide of the invention may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a protein or peptide of the invention and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3: 263-281) while the enterokinase cleavage site provides a means for purifying the protein or peptide of the invention from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; DNA Cell Biol. 12:441-453).

ACCase activity can be measured spectrophotometrically and also through the isotope exchange technique. ACCase activity was measured spectrophotometrically by coupling the production of ADP to the oxidation of NADH using pyruvate kinase and lactate dehydrogenase. This assay was used to measure overall ACCase activity by supplying acetyl CoA as a substrate. Activity of the full length Ustilago ACCase is detailed in **Figure 2**. ACCase activity was measured by way of isotope exchange based upon the fact that ACCase catalyzes the formation of malonyl CoA from acetyl-CoA and bicarbonate. The isotope exchange assay is designed to monitor the incorporation of ^{14}C from bicarbonate into the malonyl CoA product.

Binding assays were conducted to detect binding to the BC domain, and thus, enabled identification of test compounds. Methods of conducting binding assays are well known in the art. Direct measurement of the binding of radiolabeled ligands is typically performed by incubating numerous concentrations of radioligand with a constant amount of target peptide under equilibrium binding conditions followed by determining the amount of labeled probe specifically bound. "Specifically bound" is defined as total binding minus non-specific binding, where non-specific binding is determined in the presence of excess unlabled ligand. The strength of the binding interaction between the BC domain and soraphen is high and comparable to binding of soraphen to the full length ACCase (**Figures 2, 3** and **4**). The dissociation constant of the BC domain complexes, for example for soraphen, is from about 10^{-4} to

about 10^{-14} M, and preferably at 10^{-7} M to 10^{-14} M, and more preferably at least 10^{-8} M to 10^{-14} M, and still more preferably lower than 2 x 10^{-9} M. Complexes can be formed by covalent or noncovalent interactions. Once one possesses a radiolabeled ligand that binds to a target protein, any additional compound that is not radiolabeled can be assayed for binding to the same site using a competition binding assay. A competition binding assay is performed by incubating a constant concentration of target protein and radiolabled ligand with numerous concentrations of test compounds under equilibrium conditions followed by determining the amount of radiolabeled ligand that is specifically bound.

Thus, a further aspect of the present invention is a composition comprising: *(a)* an aqueous carrier solution; and *(b)* the peptide (or ACCase BC domain described herein) solubilized in said aqueous carrier solution. The composition is useful for, among other things, the binding or screening assays described herein. In general, "solubilized" means that the peptide is homogeneously or uniformly dissolved or dispersed in the carrier solution in a manner that makes the peptide in the composition available for participation in the binding events (*e.g.*, soraphen binding) described herein. The carrier solution may be any suitable aqueous solution that comprises, consists of or consists essentially of water, along with other typical optional ingredients such as buffers, agents for adjusting pH, preservatives, etc. In general, the peptide is included in the composition in any suitable amount, for example from 0.001, 0.01 or 0.1 nanograms up to 0.1, 1, 10, or 20 milligrams per milliliter of aqueous carrier solution. The peptide is in a physical form in the composition that renders it suitable for a binding assay and thus has a soraphen dissociation constant in said composition of, for example, from 10^{-4} up to 10^{-14} M. The pH of the composition may be at, or adjusted to be at, a pH suitable for binding studies, such as a pH of 5 through 9. Preferably, the Evolutionary Chemistry process as referenced herein could be utilized to identify test compounds that bind to the BC domain. In such instance, the composition should be comprised of an aqueous carrier solution containing a BC domain peptide, possessing a soraphen dissociation constant of 10^{-8} to 10^{-9} M. The BC domain would be utilized at a concentration of 0.02 to 20 milligrams per milliliter at pH 7 incubated in combination with one or more RNA-tethered test compounds for 1 hour to enable equilibrium binding to occur. Depending upon the level of stringency applied, low affinity test compounds would be washed away from the BC

domain peptide and high affinity binding compounds would be retained. The retained compounds are potential ACCase inhibitors.

Alternatively, ACCase inhibitors could be identified in a screen based on the principle of competition binding with soraphen. As mentioned previously, one way to detect competitive binding is by use of radiolabeled soraphen. In such instance, the composition should be comprised of an aqueous carrier solution containing a BC domain peptide, possessing a soraphen dissociation constant of 10^{-8} to 10^{-9} M. The BC domain would be utilized at a concentration of 5nM to 10nM at pH 7 incubated in combination with ^{3}H-soraphen (with a specific activity greater than 500 cpm per picomole) at a concentration 10% to 90% that of the BC domain, and with one or more test compounds at a concentration of 10^{-4} M to 10^{-10} M for 1 hour or more to enable equilibrium binding. The amount of ^{3}H-soraphen that remains bound to the BC domain would then be determined. A reduction in the amount of bound soraphen indicates that the test compound or compounds can bind to the same site on the BC domain and thus represent potential ACCase inhibitors.

Another preferred method to screen for ACCase inhibitors based on the principle of competitive binding with soraphen is by a fluorescence polarization assay. In this method, a fluorescent soraphen derivative that retained high affinity for BC domains would need to be acquired or prepared through standard synthetic procedures. In such instance, the composition should be comprised of an aqueous carrier solution containing a BC domain peptide, possessing a fluorescent-soraphen derivative dissociation constant of 10^{-6} M to 10^{-9} M. The BC domain would be utilized at a concentration approximately equal to the dissociation constant of the fluorescent probe, and incubated in the presence of one or more test compounds at a concentration of 10^{-4} M to 10^{-10} M for 1 hour or more to enable equilibrium binding. The fluorescence polarization would then be measured. Since the fluorescence polarization is directly related to the amount of fluorescent-soraphen derivative bound, a reduction in fluorescence polarization indicates that the test compound or compounds can bind to the same site on the BC domain and thus represent potential ACCase inhibitors.

Evolutionary chemistry (EC) as described herein relates to the process wherein product libraries are formed by combining a pool of first chemical reactants coupled to a nucleic acid with a pool of free chemical reactants. The coupled nucleic acid is

capable of mediating the chemical reaction which leads to the product library and further the nucleic acid is amplifiable so a product which has a predetermined desirable characteristic can be enriched for and identified from the product library. In its most general form, a nucleic acid-reactant test mixture is formed by attaching a first reactant (R) to each of the nucleic acids in a test mixture (containing 10^2 to 10^{18} nucleic acids with randomized sequences). The nucleic acid-reactant test mixture is treated with other free reactants that will combine with the first reactant (R) to form different products. It is important to note that from the nucleic acid test mixture (NA), discrete nucleic acid sequences will be associated with facilitating the formation of the different shaped products and are denoted, for example, by sequence-A, sequence-B and sequence-C. The products may differ in shape, reactivity or both shape and reactivity. Partitioning of the desirable product shape or reactivity is accomplished by binding to or reaction with a target. Proteins, small molecules, lipids, saccharides, etc., are all examples of targets (T). After binding to or reacting with the target the non-interacting products, which are attached to sequence-B and sequence-C are separated from sequence-A and discarded. The nucleic acid sequence-A is then amplified by a variety of methods known to those experienced in the art. Sequence-A is then used to facilitate the assembly of the desirable product by facilitating the specific reaction to form the selected product on treatment with the mixture of starting reactants. In a typical reaction, Sequence-A can be reattached to the first reactant, however, said reattachment is not always required. This is an idealized case and in many examples the nucleic acid facilitator may assemble more than one product from the starting mixture, but all of the products selected will have the desired properties of binding to or chemical reaction with the target. EC is more fully described in U.S. Patent Nos. 6,048,698; 6,030,776; 5,858,660; 5,789,160; 5723,592; and 5,723,289.

In sum, BC peptide domains, as exemplified by Ustilago pCS8, are expressed at high levels, can be purified to homogeneity, are stable under typical laboratory conditions, and exhibit high affinity soraphen binding comparable to that of full length ACCase (**Figure 5**). Therefore, it is an excellent agent for use in target based affinity binding screens and selections, including but not limited to evolutionary chemistry selections, for the identification of ACCase inhibitors.

In some assays it may be desirable to use a first peptide of the present invention in conjunction (*e.g.*, sequentially or simultaneously) with a second peptide that serves as a counterselection agent. For one embodiment, the counterselection agent may be a peptide of the same species as the first peptide (that is, with substantially the same amino acid sequence as the first peptide), but with a nonfunctional BC domain (for example, by introduction of a deletion or substitution mutation therein), to select against agents that bind non-specifically (*e.g.*, not at the soraphen binding site) to the first peptide. An example would be an *S. cerivasae* first peptide and a corresponding *S. cerivasae* second peptide in which the second peptide contains a mutation that disrupts soraphen binding (*e.g.*, S77->Y). In another embodiment, the second peptide counterselection agent may be a peptide of a different species as the first peptide, but with a functional BC domain, to detect agents that bind to and act on the first species but not the second species. For example, the first peptide may be non-mammalian, and the second peptide may be ammalian or human (*e.g.*, to select against agents that are active on the mammalian or human ACCase). Where a species contains two different ACCases such as does human, the first and second peptide may be of the same species but a different ACCase (*e.g.*, human ACC1 and human ACC2). In either embodiment, the first and second peptides can be provided together as kits or sets, either *per se* or as compositions/formulations as described above, which may be stored, utilized and/or packaged together, optionally including instructions for their use in assays as described herein.

While the present invention has been described primarily with reference to a ACCase BC domains isolated from *Ustilago maydis* (**Figure 6**; **SEQ ID NO: 2**), *Phytophthora infestans* (**Figure 7**; **SEQ ID NO: 4**), *Magnaporthe grisea* (**SEQ ID NO: 6**), *Saccharomyces cerevisiae* (**SEQ ID NO: 8**), *and Homo sapiens* (**SEQ ID NOS: 10, 12, 14, 16**), it will be appreciated that distantly related organisms may be substituted for the organisms described herein. For example, peptides of the present invention may be isolated from other fungal, insect, or plant species, such as set forth in **Tables 1-3** below, including any other members of the kingdoms, divisions, classes, orders or families set forth therein, as well as nematodes and mammals.

Table 1. Fungal Pests (Kingdom = Fungi if Division is not Oomycota; If Division = Oomycota, then Kingdom = Chromista)

Genus	Species	Common Name	Family	Order	Class	Division	Major Crop
Magnaporthe	grisea	rice blast	Magnaporthaceae	Diaporthales	Ascomycetes	Ascomycota	rice
Erysiphe	graminis tritici	powdery mildew	Erysiphaceae	Erysiphales	Ascomycetes	Ascomycota	wheat
Septoria (Leptosphaeria)	nodorum and tritici	septoria	Leptosphaeriaceae	Pleosporales	Ascomycetes	Ascomycota	wheat
Gaeumannomyces	graminis	take-all	Pythiaceae	Pythiales	Oomycetes	Oomycota	wheat
Pythium spp			Pythiaceae	Pythiales	Oomycetes	Oomycota	turf
Puccinia	sorghi	stalk rot/rust	Pucciniaceae	Uredinales	Urediniomycetes	Basidiomycota	maize
Aspergillus	flavus		Trichocomaceae	Eurotiales	Ascomycetes	Ascomycota	maize
Phytophthora	Infestans	late blight	Pythiaceae	Pythiales	Oomycetes	Oomycota	potatoes
Fusarium spp		wilt	Nectriaceae	Hypocreales	Ascomycetes	Ascomycota	potatoes
Botrytis spp			Clavicipitaceae	Hypocreales	Ascomycetes	Ascomycota	tree/vines
Alternaria spp			Pleosporaceae	Pleosporales	Ascomycetes	Ascomycota	
Cercospora spp			Mycosphaerellaceae	Mycosphaerellales	Ascomycetes	Ascomycota	
Rhizoctonia spp			Platygloeaceae	Platygloeales	Ustilaginomycetes	Basidiomycota	
			Ceratobasidiaceae	Ceratobasidiales	Basidiomycetes	Basidiomycota	
Peronospora spp		Downy mildew	Peronosporaceae	Peronosporales	Oomycetes	Oomycota	
Colletotrichum spp			Glomerella	Glomerellaceae	Ascomycetes	Ascomycota	
Bremia spp		Downy mildew	Peronosporaceae	Peronosporales	Oomycetes	Oomycota	

Table 2: Insect Pests (Kingdom = Animalia; Phylum = Arthropoda)

Genus	Species	Common Name	Family	Order	Class	Major Crop
Nilaparvata	lugens	Brown planthopper	Delphacidae	Hemiptera	Insecta	rice
Mayetiola	destructor	Hessian fly	Cecidomyiidae	Diptera	Insecta	wheat
Heliothis	zea	Corn earworm/ bollworm	Noctuidae	Lepidoptera	Insecta	maize
Ostrinia	nubilalis	European cornborer	Pyralidae/Crambidae	Lepidoptera	Insecta	maize
Diabrotica spp		Corn rootworm	Chrysomelidae	Coleoptera	Insecta	maize
Myzus spp		aphid	Aphididae	Homoptera	Insecta	potato
Leptinotarsa	decemlineata	Colorado beetle	Chrysomelidae	Coleoptera	Insecta	potato
Pectinophora	gossypiella	Pink bollworm	Gelechiidae	Lepidoptera	Insecta	cotton
Heliothis spp			Noctuidae	Lepidoptera	Insecta	cotton
		whiteflies	Aleyrodidae	Homoptera	Insecta	
		Potato leafhopper	Cicadellidae	Hemiptera	Insecta	
Plutella	xylostella	Diamondback moth	Plutellidae	Lepidoptera	Insecta	
Chaetocnema spp		Flea beetles	Chrysomelidae	Coleoptera	Insecta	

5

Table 3: Weedy Pests (Kingdom = Plantae; Division = Magnoliophyta)

Genus	Species	Common Name	Family	Order	Class	Major Crop
Echinochloa	crus-galli	Barnyard grass	Poaceae	Cyperales	Liliopsida	rice, cotton
Echinochloa	colonum		Poaceae	Cyperales	Liliopsida	maize
Avena	fatua	Wild oats	Poaceae	Cyperales	Liliopsida	wheat
Polygonum	convolvulus	black bindweed	Polygonaceae	Polygonales	Magnoliopsida	wheat
Cyperus	rotundus	sedge	Cyperaceae	Cyperales	Liliopsida	maize, cotton
Chenopodium	album	lambsquarters	Chenopodiaceae	Caryophyllales	Magnoliopsida	potato
Galium spp		bedstraw	Rubiaceae			
Ipomoea spp		morningglory	Convolvulaceae	Solanales	Magnoliopsida	
Amaranthus spp		pigweed	Amaranthaceae	Caryophyllales	Magnoliopsida	
Digitaria spp		Crabgrass	Poaceae	Cyperales	Liliopsida	
Lolium spp		ryegrass	Poaceae	Cyperales	Liliopsida	
Sorghum	halepense	Johnson grass	Poaceae	Cyperales	Liliopsida	
Panicum	miliaceum	Wild proso millet	Poaceae	Cyperales	Liliopsida	
Senna spp		Velvet leaf	Fabaceae	Fabales	Magnoliopsida	

The present invention is explained in greater detail in the following non-limiting Examples and the Figures herein, in which the following abbreviations are used: pCS8 - **Ustilago maydis**, basidomycete, N-terminal His-tag BC domain, 64.6 kDa protein; pCS11- **Ustilago maydis**, basidomycete, full length ACCase with C-terminal His-tag, 241.4 kDa protein; pCS15 - **Phytophthora infestans**, oomycete C-terminal His-tag BC domain, 63.3 kDa protein; pCS16 - **Saccharomyces cerevisiae,** wild type N-terminal His-tag BC domain, 68.0 kDa protein; pCS16M - **Saccharomyces cerevisiae,** S77Y mutant, N-terminal His-tag BC domain, 68.0 kDa protein; pCS17 - **Magnaporthe grisea**, ascomycete, N-terminal His-tag BC domain, 68.2 kDa protein; pCS19 - **Homo sapiens**, C-term His-tag ACC1 BC domain, 71.2 kDa protein; pCS20 – **Homo sapiens**, C-term His-tag ACC2 BC domain, 86.7 kDa protein; pCS201 – **Escherichia coli**, N-terminal His-tag BC protein, 51.6 kDa protein; pCS204 - **Saccharomyces cerevisiae,** C-terminal His-tag wild type full length ACCase protein, 254.3 kDa protein; and pCS204M - **Saccharomyces cerevisiae,** C-terminal His-tag S77Y mutant full length ACCase protein, 254.3 kDa protein.

EXAMPLE 1

A. Preparation of BC Domain Peptide. *E.coli* cultures transformed with protein expression constructs for BC domains with either N or C terminal his-tags (as illustrated in **Figure 5** for pCS8) were induced by the addition of IPTG (0.2 mM) at an OD_{600} = 0.5. The cultures were grown overnight at 18 °C, harvested and stored at –80°C. The bacterial pellet was resuspended in a buffer containing 50 mM NaH_2PO_4 (pH8), 300 mM NaCl, 10 mM imidazole, protease inhibitor and 1 mg/mL lysozyme. The lysate was sonicated and nuclease was added. The lysate was then incubated with Ni-NTA resin (Novagen) for 1 hour at 4°C. pCS8 was eluted with a buffer containing 50 mM NaH_2PO_4 (pH8), 300 mM NaCl, and 250 mM imidazole. Fractions containing pCS8 were combined and ammonium sulfate precipitated (40% ammonium sulfate). The pellet from the ammonium sulfate precipitation was resuspended in SB (SB = 200mM NaH_2PO_4 pH 7.0, 10% glycerol). Protein concentrations were determined by Bradford analysis. The purified protein was stored at –80 °C.

B. Purification of BC Domain Peptide. A one-step purification of his-tagged BC domain peptides, as exemplified for pCS8, on Ni-NTA-agarose yields protein that is approximately 90-95% pure as judged by SDS-PAGE was utilized. This method is similar to the purification performed with a histidine-tag attached to the amino terminus of a mutant form of the enzyme and nickel affinity chromatography as described in C. Blanchard et al., "Mutations at Four Active Site Residues of Biotin Carboxylase Abolish Substrate-Induced Synergism by Biotin," Biochemistry, vol. 38, pp. 3393-3400 (1999). After elution from the Ni-NTA agarose column, the pCS8 protein was precipitated by adding an equal volume of saturated ammonium sulfate. The precipitated protein was then resuspended in SB to a concentration of 10 to 20 milligrams per milliliter and stored at –80 °C until used. This method is utilized for purification of all BC domains described herein. Because purer preparations may be required for some purposes, an additional polishing step was investigated. For this purpose, a single UNO-Q (Bio-Rad) anion exchange step subsequent to the Ni-NTA-agarose chromatography purified pCS8 to apparent homogeneity with good yield. UNO-Q is a fast flow matrix that is readily amenable to scale-up. *See also* **Figure 8.**

EXAMPLE 2

ACCase Activity Assays. The following methods were employed to detect ACCase activity. *See also* **Figure 9A and 9B.**

Method 1: Assay-Spectrophotometric. ACCase activity was measured spectrophotometrically by coupling the production of ADP to the oxidation of NADH using pyruvate kinase and lactate dehydrogenase. This assay was used to measure either overall ACCase activity by supplying acetyl CoA as a substrate, or BC activity by supplying free biotin as a substrate (note, however, that this has only been demonstrated with "prokaryotic type" BC's). This assay is best suited for purified protein. This assay would be used to test for enzymatic activity of compounds identified by virtue of their binding to isolated BC domains, including but not limited to the EC process.

To establish this assay, *E. coli* biotin carboxylase was first cloned and expressed according to the literature (Biochemistry 38:3393-3400, 1999). As seen in **Figure 9A**, the activity of *E coli* BC towards a free biotin substrate was readily detectable. Under the conditions of the assay, activity was detected with as little as 40

ng protein, and a maximal velocity was reached between 2 and 20 µg pCS201. In repeated attempts with multiple pCS8 preparations, however, no activity was detected using up to 7 µg protein. We conclude that pCS8 is unable to carboxylate free biotin, as would be expected since such activity has not been detected utilizing eukaryotic BC domains.

To measure overall ACCase activity, a full-length Ustilago maydis ACCase was cloned, expressed in E. coli with a C-terminal His-tag (pCS11), and purified. As seen in Figure 9B(i), the time dependent oxidation of NADH, detected by a decrease in absorbance at 340nm, was dependent on both pCS11 protein as well as acetyl CoA substrate.

Method 2: Assay-^{14}C Isotope Exchange. ACCase catalyzes the formation of malonyl CoA from acetyl-CoA and bicarbonate. The isotope exchange assay is designed to monitor the incorporation of ^{14}C from bicarbonate into the malonyl CoA product. Malonyl CoA is acid and heat stable so the unreacted $H^{14}CO_3^-$ can be removed by acidification followed by evaporation. As can be seen in Figure 9B(ii) and consistent with the spectrophotometric assay, activity (measured as the incorporation of ^{14}C into malonyl CoA) was dependent on both pCS11 protein and acetyl CoA substrate. Because this an endpoint assay, it is less suitable than the spectrophotometric assay for kinetic measurements; however, it is superior for detecting activity in crude preparations. It will also be used to test for enzymatic activity of compounds identified by virtue of their binding to isolated BC domains, including but not limited to the EC process.

EXAMPLE 3

Soraphen binding assay. Affinity based screens and selection assays, including but not limited to EC selections, rely on binding of small molecules and not inhibition of enzymatic activity. Therefore, despite the lack of enzymatic activity, pCS8 would be a suitable affinity based screening or selection agent if it retained the high affinity soraphen binding activity of the full-length *Ustilago* ACCase. Soraphen A was tritium labeled by Sibtech, Inc. (Newington, CT) and used for binding experiments with pCS8 protein. Briefly, 6.7 nM purified pCS8 protein was incubated with various concentrations (approximately 0.5-20 nM) of ^{3}H-soraphen in PNT buffer (100mM NaH_2PO_4, 150 mM NaCl, 0.01% Triton X-100, pH 7.0) for 45 min at room

temperature. pCS8 protein (with bound ligand) was separated from free ligand on NAP-5 desalting columns (Amersham Biosciences) and the amount of bound ^{3}H-soraphen determined by liquid scintillation counting. Non-specific binding was determined with duplicate samples containing 2 μM cold soraphen. The data were fit by non-linear regression to a one site ligand binding equation: $Y=B_{max}*X/(K_D+X)$ where Y=bound ligand and X=free ligand.

pCS8 exhibited saturable binding of ^{3}H-Soraphen consistent with a single high affinity binding site (**Figure 3**). The data shown for pCS8 are combined from two experiments with independent protein preps demonstrating little prep-to-prep variability. A negative control was provided by pCS201. pCS201 encodes an N-terminal His-tagged enzymatically active *E.coli* BC that is not inhibited by soraphen. As expected, pCS201 did not exhibit high affinity soraphen binding.

Non-linear regression fits of the data gave an estimate of 1.5 nM for the K_D of the soraphen-pCS8 interaction (**Figure 4**). This is in good agreement with the K_D estimate of the soraphen-pCS11 full length ACCase of 1.6 nM (**Figure 2**) and is in good agreement with the published value of a 1.4 nM K_i for soraphen inhibition of *Ustilago* ACCase activity (Heike Behrbohm Ph.D. thesis, Braunschweig Techn. Univ., 1996). As such, the pCS8 is a suitable affinity-based screening and selection agent as it retains high-affinity soraphen binding comparable to full-length *Ustilago* ACCase.

EXAMPLE 4

Additional characterization of pCS8. No protein degradation or loss of soraphen binding was seen after incubation of SB solubilized pCS8 peptide for 24 h at room temperature. No protein degradation or loss of soraphen binding was seen after storage of SB solubilized pCS8 peptide for 5 weeks at -80 °C, including multiple freeze thaws.

EXAMPLE 5

Partial summary of BC domains generated. A number of biotin carboxylase (BC) domains that have been characterized herein and can be purified in sufficient quantities for use in affinity based screens or selections , including but not limited to selections using Evolutionary Chemistry, and five BC domains are as follows: wild

type versions from *Ustilago maydis*, *Phytophthora infestans*, *Magnaporthe grisea*, and *Saccharomyces cerevisiae*; and the mutated version from *Saccharomyces cerevisiae*. Amino acid sequence alignments of the expressed domains are shown in **Figure 10**. Identical residues are indicated with an asterisk, and the S to Y mutation in *S. ceravisae* that abolishes soraphen binding is indicated in bold.

EXAMPLE 6

Generation of the *Magnaporthe grisea* BC domain. A preliminary sequence of the entire genome of *Magnaporthe grisea* (ascomycete, causal agent of rice blast) was recently released into the public domain by the Whitehead Institute in collaboration with Ralph Dean's lab at North Carolina State University. The full-length *M. grisea* ACCase gene was PCR amplified from genomic DNA, cloned, and sequenced. One small predicted intron was removed to create a full-length cDNA. The biotin carboxylase domain was subcloned (based on alignment with our pCS8 *U. maydis* BC domain) and inserted into a pET vector (5'His tag) to make pCS17. pCS17 was expressed in E. coli and the His-tagged BC domain was purified and assayed for soraphen binding. As expected, the Magnaporthe BC domain exhibited high affinity soraphen binding, as demonstrated in **Figure 11**.

EXAMPLE 7

Cloning and expression of human ACCase genes. There are at least two forms of the acetyl-CoA carboxylase enzyme in humans. ACC1 is a cytosolic enzyme present at high levels in liver and lipogenic tissues, and is the primary species responsible for fatty acid synthesis. ACC2 is a mitochondrial enzyme found primarily in heart and muscle tissue, and is thought to regulate fatty acid oxidation. Biotin carboxylase domains from human ACCases could potentially be useful as counterselection agents with potential to select against mammalian toxicity. Additionally, agonists and inhibitors of human ACCase that can distinguish between ACC1 and ACC2 BC domains may have potential pharmaceutical applications.

Complete genomic DNA sequences are available for both genes. Their BC domains (based on homology with our pCS8 ustilago clone) were successfully cloned by amplifying small exons from genomic DNA and using PCR to splice them together. The ACC1 BC domain consists of 14 exons, 104bp to 237 bp in length,

which were assembled to make an 1896 bp (**SEQ ID NO: 9**) BC domain cDNA. The ACC2 BC domain consists of 14 exons, 108-661 bp in length, assembled to make a 2322 bp (**SEQ ID NO: 11**) BC domain cDNA. Both BC domains were cloned into pET30 to make 3' His-tagged fusion proteins. The ACC1 clone is designated pCS19 and should produce a fusion protein of 71.2kD. The ACC2 clone is designated pCS20 and should produce a fusion protein of 86.7kD. Expression analysis showed that the ACC1 fusion protein is expressed at low levels in E. coli and can be purified from the soluble fraction. The purified pCS19 protein exhibited high affinity soraphen binding (see **Figure 12**). Alignments of the Ustilago and human ACCase domains are depicted in **Figure 13**. Besides being useful in identifying selective agrochemicals, a particularly intriguing use of the human BC domains is to identify specific inhibitors that preferentially target the ACC2 domain, but not the ACC1 domain, since such inhibitors could prove useful in controlling body weight.

EXAMPLE 8

This example demonstrates that pCS8has binding characteristics amenable for use in affinity based screening or selection procedures.

Soraphen off-rate determination. Kinetic aspects of binding interactions are an important parameter in designing optimal conditions for affinity-based screening and selection assays, including but not limited to EC. Therefore, soraphen dissociation experiments were performed to determine the off-rate for the binding interaction using the *Ustilago* BC domain (pCS8). 53.6 pmol pCS8 protein was incubated with 30 pmol ^{3}H-soraphen in a volume of 2 ml for 15 min. The protein with bound soraphen was separated from free soraphen on NAP5 columns and eluted in a total of 4 ml. Cold soraphen was added to 2 μM and dissociation of the bound soraphen was followed by removing 0.5 ml aliquots at various times and applying them to NAP 5 columns to again separate bound from free radiolabel. The ^{3}H -soraphen in the eluted fractions was quantified by liquid scintillation counting and the data were fit to the following equation: $Y=Y_{max}*e^{-kt} + NS$ (**Figure 14**).

This off rate corresponds to 10.7 min half-life. Note that the on-rate can be calculated from this data since $K_d=k_{off}/k_{on}$; so $k_{on} = 9.31 \times 10^5\ M^{-1}\ s^{-1}$. At 10 μM protein, even with very low ligand concentrations, such a binding interaction would approach equilibrium very rapidly.

EXAMPLE 9

Competitive Binding Assay. Isolated biotin carboxylase domains and radiolabeled soraphen can be employed in a competitive binding assay to test the ability of any compound to bind to the soraphen binding site. Like soraphen, such compounds are likely to inhibit ACCase activity. To exemplify this assay, we prepared two soraphen derivatives with modifications at the 5-position (soraphen A-conjugate) or 11-position (soraphen C-conjugate). Both derivatives were then tested in competition binding assays with both pCS8 and pCS11 proteins. The proteins were incubated with ^{3}H-soraphen and various concentrations of the 2 conjugates for 1 hr. Bound ^{3}H -soraphen was then separated from unbound on NAP 5 columns and quantified by scintillation counting. To estimate the conjugates K_i's, the resulting data was fit to an equation for heterologous competitive binding with ligand depletion (H.J. Motulsky, Analyzing Data with GraphPad Prism, 1999, GraphPad Software Inc., San Diego CA, www.graphpad.com):

$$Y = \frac{[free]*Bmax}{[free] + Kd\,(1 + [cold]/Ki)} + NS$$

Representative data are shown in **Figure 15**, in which the soraphen A-conjugate is labeled A-conj and the soraphen C-conjugate is labeled C-conj. As expected, there was no significant difference between the affinities of the conjugates for the BC domain (pCS8) (**Figure 15A**) and the full-length protein (pCS11) (**Figure 15B**). As a positive control for this assay, an experiment using cold soraphen A as the competitor was performed and yielded a K_i estimate of 1.1nM (data not shown), consistent with direct saturation binding experiments.

EXAMPLE 10

Soraphen resistant *Saccharomyces cerevisiae* ACCase and ACCase BC domain mutants – Counterselection Agents. Mutation of Serine 77 to Tyrosine in the S. cerevisiae ACCase protein has been shown to confer soraphen resistance (European Patent Application 94810710, 1994; and United States Patent 5,641,666, 1997).

This soraphen resistance mutation was introduced into pCS204, a yeast expression vector containing the full-length *S. cerevisiae* ACCase gene constructed by cloning a PCR-amplified full-length S. *cerevisiae* ACCase gene into the expression vector pYES2 for inducible overexpression of His-tagged ACCase in *S. cerevisiae*. The resulting construct was designated pCS204M. The *S. cerevisiae* biotin carboxylase domains from pCS204 and pCS204M were then subcloned into an *E.coli* pET expression vector to form pCS16 and pCS16M, respectively. Like pCS8, both of these constructs express their respective BC domains as N-terminal-His-tagged fusion proteins to facilitate purification. Expression analysis demonstrated that both pCS16 and pCS16M yielded comparable amounts of purified, soluble protein. The products were then analyzed by size exclusion chromatography and, like pCS8 protein, both were found to exist primarily as monomers (>90%). The proteins were then tested for soraphen binding and the results are shown in **Figure 16**. pCS16 protein exhibited high-affinity soraphen binding comparable to pCS8 protein. In contrast, soraphen binding by pCS16M was similar to the non-specific control (Data not shown) confirming that introduction of this single amino acid mutation into a BC domain abolishes soraphen binding. Therefore, pCS16M protein would be an excellent counter-selection agent to eliminate non-soraphen-binding-site interactions.

The effect of the mutation on full-length ACCase was also assessed. pCS204 and pCS204M were overexpressed in *S. cerevisiae* and purified by Ni-NTA chromatography. Both proteins appeared identical on SDS-PAGE (Data not shown).

The proteins were then assayed for ACCase activity using the ^{14}C isotope exchange assay. The resulting data were similar to those from a published report (*Curr. Genet.* **25**:95-100 (1994)) comparing the activity of the endogenous protein from wild type and mutant yeast, and demonstrate that pCS204M protein activity is insensitive to soraphen but still sensitive to avidin inhibition (**Table 4**).

Table 4

	Relative Enzymatic Activity (%)			
Treatment	WT enzyme*	Mutant enzyme*	pCS204	pCS204M
Control	100.0	100.0	100.0	100.0
+150 µg/mL soraphen A	0.8	74.7	7.5	105
+1.5 µg/mL soraphen A	1.0	74.9	11.6	101
+1.5 ng/mL soraphen A	71.8	79.4	92.0	103
+250 µg/ml avidin	0.4	0.1	2.1	2.5
- acetyl CoA	0.6	0.2	2.0	2.8

*Data from *Curr. Genet.* **25:**95–100 (1994).

Finally, pCS204 and pCS204M were assayed for soraphen A binding and the results are shown in **Figure 17**. Full-length pCS204 protein bound soraphen with similar high affinity as the BC domain expressed by pCS16. Soraphen binding by pCS204M, like that of pCS16M, was comparable to the non-specific control.

The foregoing is illustrative of the present invention, and is not to be construed as limiting thereof. The invention is defined by the following claims, with equivalents of the claims to be included therein.